Shift-Share IV Design

Prof. Tzu-Ting Yang
楊子霆

Institute of Economics, Academia Sinica
中央研究院經濟研究所

May 13, 2026

Main Idea

Overview

- The Shift-Share Instrumental Variable (SSIV) is a commonly used IV approach in econometrics to identify causal effects
- It decomposes changes in economic variables within regions into two components:
  - **Shift**: Aggregate/external changes in economic variables across groups (e.g., industries, demographics)
  - **Share**: Local area's initial exposure or composition of these groups
- First introduced by Bartik (1991), further developed and applied by Autor et al. (2013) and Kovak (2013)
  - Commonly applied in labor economics, trade, and economic geography

Motivating Example

Impact of Robot on Labor Market Outcomes

- We want to examine the effects of automation/robot on regional labor market by estimating the following regression

$$\Delta Y_{it} = \kappa + \beta \Delta R_{it} + X_{it-\tau}\gamma + u_{it}$$

- ΔY_{it}: Change in employment in the district i over a long period
  - Δ represents long-run differences (e.g. 10 years, $\tau = 10$)
  - $\Delta Y_{it} = Y_{it} - Y_{it-\tau}$
  - Y_{it} represents total employment or average wage

$$\Delta Y_{it} = \kappa + \beta \Delta R_{it} + X_{it-\tau}\gamma + u_{it}$$

- ΔR_{it}: Change in robot exposure measure in the district i over a long period
  - $\Delta R_{it} = R_{it} - R_{it-\tau}$
  - R_{it} represents robot stock
- $X_{it-\tau}$: Control variables at district level in the baseline period $t - \tau$

$$\Delta Y_{it} = \kappa + \beta \Delta R_{it} + X_{it-\tau}\gamma + u_{it}$$

- β: Impact of robot exposure on labor market outcomes
- **Challenge:** Isolating the effect of robot exposure from other regional economic factors
  - Observed robot stock is endogenous to labor market conditions
  - The increases in product demand for the factories in Hsinchu city would increase both the employment/wage and robot stock of the area

Deriving the Shift-Share IV

- We can decompose the endogenous variable, ΔR_{it}, as a weighted average of industry-specific changes:

$$\Delta R_{it} \approx \sum_{k=1}^{N} \overbrace{\Delta R_{ikt}}^{\text{shift}} \times \overbrace{\frac{L_{ikt}}{L_{it}}}^{\text{share}}$$
$$\approx \Delta R_{i1t} \times \frac{L_{i1t}}{L_{it}} + \Delta R_{i2t} \times \frac{L_{i2t}}{L_{it}} + \cdots + \Delta R_{iNt} \times \frac{L_{iNt}}{L_{it}}$$

  - $k = 1, 2, \ldots, N$ indexes the industries in the economy
  - $\frac{L_{ikt}}{L_{it}}$ is industry k's employment share in region i at time t
  - ΔR_{ikt} is the change in outcome for industry k in region i
  - This is an approximation, not an identity. It holds closely only if employment shares are relatively stable between $t - \tau$ and t

Derive Shift-Share IV

Step 1

Step 1: Delocalization Shift IV

- Replace the original **shift part**, ΔR_{ikt}, by the **global/national change** ΔR_{kt}
  - The national change in robot stock for a industry k

$$\Delta Z_{it} = \sum_{k=1}^{N} \overbrace{\Delta R_{kt}}^{shift} \times \overbrace{\frac{L_{ikt}}{L_{it}}}^{share}$$

- Sometimes ΔR_{kt} may be $\Delta R_{kt,-i}$. It's the global change calculated without district i's values.

Step 2: Shift-Share IV

- We add one more revision to the time period t of the **share part**:
  - Replace it by some rules. E.g., frozen Lag (use base period's share), updating lag, averaging all periods.

$$\Delta Z_{it} = \sum_k \overbrace{\Delta R_{kt}}^{shift} \times \overbrace{\frac{L_{ikt-\tau}}{L_{it-\tau}}}^{share}$$

Why Shift-Share IV Wokrs?

Intuition

- **De-localizing shift** should be able to remove the main part that causes selection bias
  - Use the national/global changes that cannot be affected (at least not too much) by the local factors to replace the shift.
- But, the **share part** may be affected by confounders.
  - Therefore, we may use **Frozen Lag** (a base period's share), an **Updating Share** (use t-k share), or **averaging share**
  - Make sure the IV is **totally prior to the current period**
- Intuitively, we want the **share part** to reflect that district's geographical or historical features that cannot be easily affected

- **Method 1: Frozen Lag**
  - Construct share using the value in first period $t = 0$
  - So the first period of the panel is dropped
- Downside:
  - Because the first period t=0 has a different distance from each t in the panel, the later periods are likely to be less relevant to $t = 0$ time period.

- **Method 2: Updating Lag**
  - Construct share using the value in period $t - \tau$
  - Update share every τ periods
  - τ is a fixed number

- **Method 3: Averaging over all periods**
  - Construct share using the average value during all periods
  - With a large number of time periods, that is, the role of any one time period's values is minor

Why Shift-Share IV Works?

$$\Delta Z_{it} = \sum_{k=1}^{N} \overbrace{\Delta R_{kt}}^{\text{shift}} \times \overbrace{\frac{L_{ik,t-\tau}}{L_{i,t-\tau}}}^{\text{share}}$$

- **First-stage Relevance**: National/global changes (ΔR_{kt}) are strongly correlated with local changes (ΔR_{ikt})
- **Exclusion Restriction**: For validity, the instrument should affect outcomes only through the endogenous variable
  - Historical shares ($\frac{L_{ik,t-\tau}}{L_{i,t-\tau}}$) are predetermined, reducing concerns about simultaneity
  - National shifts (ΔR_{kt}) are assumed to be exogenous to local conditions

Potential Threats to SSIV Validity

- The exclusion restriction may fail for several reasons:
  - **Persistent local shocks**: Historical industry shares may correlate with unobserved local characteristics that affect both past and current outcomes
  - **Feedback effects**: Large regions might influence national shifts, violating the exogeneity assumption (especially relevant for large metropolitan areas)
- Despite these challenges, SSIV remains valuable for addressing **reverse causality** and some forms of **omitted variable bias**

- **First Stage Regression:**

$$\Delta R_{it} = \lambda + \alpha \Delta Z_{it} + X_{it-\tau}\theta + \epsilon_{it}$$

- **Second Stage Regression:**

$$\Delta Y_{it} = \kappa + \beta \Delta \hat{R}_{it} + X_{it-\tau}\gamma + u_{it}$$

- Use shift-share as IV for treatment variable ΔR_{it}

$$\Delta Z_{it} = \sum_{k=1}^{N} \overbrace{\Delta R_{kt}}^{shift} \times \overbrace{\frac{L_{ikt-\tau}}{L_{it-\tau}}}^{share}$$

Relevant Jargons

- **A summary of relevant jargons**
  - Exposure Design $\approx$ SSIV = Bartik IV
    - Because Bartik proposed this IV
  - Shock = Shift
  - Some may say: "use predicted value to instrument the endogenous var..." since the SSIV is like a predicted value of the endogenous var.

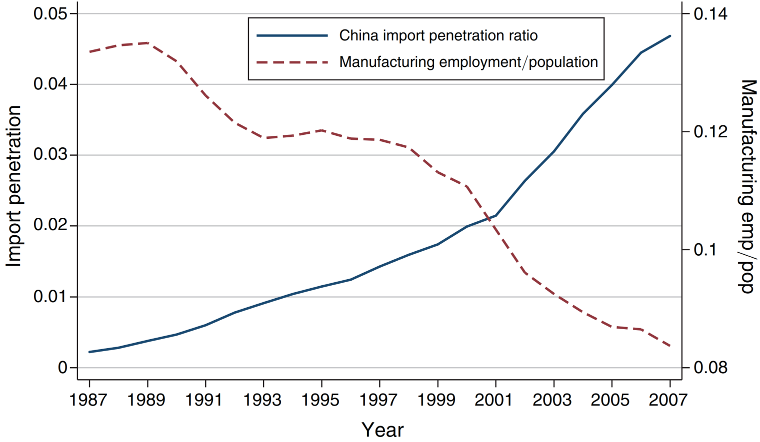
Empirical Example

Autor, David Dorn, and Gordon H. Hanson (2013)

Autor, David H., David Dorn, and Gordon H. Hanson. "**The China syndrome: Local labor market effects of import competition in the United States.**" *American Economic Review* 103.6 (2013): 2121-2168.

- The paper examines the impact of increased Chinese import competition between 1990 and 2007 on local U.S. labor markets.
- Utilizes a shift-share IV strategy to identify the effects of import competition.
- Key findings suggest significant impacts on employment, wages, and the receipt of welfare benefits in regions more exposed to import competition.

Graphical Evidence



- See SSIV_build_data.do
  - Show how to construct shift-share IV
- See SSIV.do and SSIV.R
  - Implement IV estimation
- Use location_level.dta

Shift-share IV design

- Analysis is based on cross-market variations in import exposure driven by initial industry specialization.
- Shift-share IV design:
    - Changes in Chinese imports by other high-income countries, to isolate the supply-driven component of U.S. import increases.
- Focus on manufacturing industries and their local labor market conditions across U.S. regions.

Autor, David Dorn, and Gordon H. Hanson (2013)

Step 1: Construct Treatment Variable

- They use Commuting Zone (CZ) as the observation unit
  - CZ: multiple adjacent counties that have the commuting structure of a local labor market
- The treatment variable ΔIP^{cu}_{it} measures local exposure at zone *i* to the growth of imports from China per worker (measured in \$1,000) during period t and $t-10$

$$\Delta IP^{cu}_{it} = \sum_k \overbrace{\frac{\Delta M^{cu}_{kt}}{L_{kt}}}^{shift} \overbrace{\frac{L_{ikt}}{L_{it}}}^{share}$$

  - $\frac{L_{ikt}}{L_{it}}$ the industry k's employment share of zone *i* at year t.

Autor, David Dorn, and Gordon H. Hanson (2013)

Step 1: Construct Treatment Variable

$$\Delta IP_{it}^{cu} = \sum_k \overbrace{\frac{\Delta M_{kt}^{cu}}{L_{kt}}}^{shift} \overbrace{\frac{L_{ikt}}{L_{it}}}^{share}$$

- $\frac{\Delta M_{kt}^{cu}}{L_{kt}}$ is exposure to the growth of imports from China per worker in the industry k (measured in \$1,000)
  - $\Delta M_{k\tau}^{cu}$: the growth in US imports from China during period t and $t - \tau$
  - L_{kt}: total employment in industry k at period t

The SSIV for this paper is:

$$\Delta Z_{it}^{co} = \sum_k \overbrace{\frac{\Delta M_{kt}^{co}}{L_{kt-10}}}^{shift} \overbrace{\frac{L_{ikt-10}}{L_{it-10}}}^{share}$$

- **Shift part:** $\frac{\Delta M_{kt}^{co}}{L_{kt-10}}$
  - ΔM_{kt}^{co}: the growth in other high-income countries' imports from China during period t
  - L_{kt-10}: total employment in industry k in period $t-10$ (lag 10 years)
- **Share part:** lag 10 years

TABLE 1—VALUE OF TRADE WITH CHINA FOR THE US AND OTHER SELECTED HIGH-INCOME COUNTRIES AND VALUE OF IMPORTS FROM ALL OTHER SOURCE COUNTRIES, 1991/1992–2007

	I. Trade with China (in billions 2007 US$)		II. Imports from other countries (in billions 2007 US$)		
	Imports from China (1)	Exports to China (2)	Imports from other low-inc. (3)	Imports from Mexico/CAFTA (4)	Imports from rest of world (5)
Panel A. United States					
1991/1992	26.3	10.3	7.7	38.5	322.4
2000	121.6	23.0	22.8	151.6	650.0
2007	330.0	57.4	45.4	183.0	763.1
Growth 1991–2007	1,156%	456%	491%	375%	137%
Panel B. Eight other developed countries					
1991/1992	28.2	26.6	9.2	2.8	723.6
2000	94.3	68.2	13.7	5.3	822.6
2007	262.8	196.9	31.0	11.6	1329.8
Growth 1991–2007	832%	639%	236%	316%	84%

- **First Stage Regression:**

$$\Delta IP_{it}^{cu} = \lambda + \alpha \Delta Z_{it}^{co} + X_{it-10}\theta + \epsilon_{it}$$

- **Second Stage Regression:**

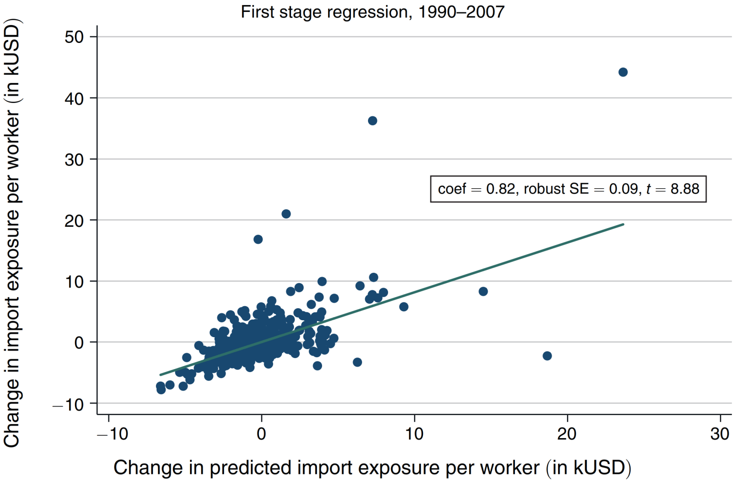
$$\Delta Y_{it} = \kappa + \beta \Delta \hat{IP}_{i\tau}^{cu} + X_{it-10}\gamma + u_{it}$$

- Use shift-share as IV for treatment variable $\Delta IP_{i\tau}^{cu}$

$$\Delta Z_{it}^{co} = \sum_k \overbrace{\frac{\Delta M_{kt}^{co}}{L_{kt-10}}}^{shift} \overbrace{\frac{L_{ikt-10}}{L_{it-10}}}^{share}$$

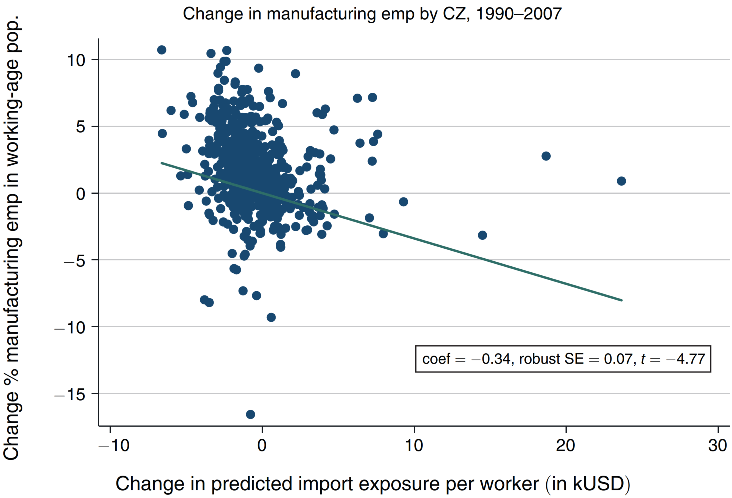
Graphical Evidence

Panel A. 2SLS first stage regression, full sample



Graphical Evidence

Panel B. OLS reduced form regression, full sample



STATA Command: ivregress

```
ivregress 2sls y (x=z)  [aw=wei], cluster(czone) first

ivregress 2sls y (x=z) l_sh_popedu_c l_sh_popfborn
    l_sh_empl_f l_sh_routine33 l_task_outsource [aw=wei],
    cluster(czone) first

ivreghdfe y (x=z) [aw=wei], cluster(czone)

ivreghdfe y (x=z) l_sh_popedu_c l_sh_popfborn l_sh_empl_f
    l_sh_routine33 l_task_outsource [aw=wei], cluster(czone)
```

- **aw=wei**: total employment at initial period
- **y**: manufacturing employment at CZ
- **x** china import exposure at CZ
- **z**: shift-share IV

R Command: ivreg

```
# Instrumental variables regression using ivreg
model3 <- ivreg(y ~ x | z, data = location_level, weights =
    wei)
summary(model3, vcov = vcovCL, cluster = ~czone)

# Instrumental variables regression with additional controls
model4 <- ivreg(y ~ x + l_sh_popedu_c + l_sh_popfborn +
    l_sh_empl_f + l_sh_routine33 + l_task_outsource | z,
data = location_level, weights = wei)
summary(model4, vcov = vcovCL, cluster = ~czone)
```

- **z**: Shift-share instrument used for IV regression

Results

TABLE 3—IMPORTS FROM CHINA AND CHANGE OF MANUFACTURING EMPLOYMENT IN CZS, 1990–2007: 2SLS ESTIMATES

Dependent variable: 10 × annual change in manufacturing emp/working-age pop (in % pts)

	I. 1990–2007 stacked first differences					
	(1)	(2)	(3)	(4)	(5)	(6)
(Δ imports from China to US)/worker	−0.746*** (0.068)	−0.610*** (0.094)	−0.538*** (0.091)	−0.508*** (0.081)	−0.562*** (0.096)	−0.596*** (0.099)
Percentage of employment in manufacturing$_{-1}$		−0.035 (0.022)	−0.052*** (0.020)	−0.061*** (0.017)	−0.056*** (0.016)	−0.040*** (0.013)
Percentage of college-educated population$_{-1}$				−0.008 (0.016)		0.013 (0.012)
Percentage of foreign-born population$_{-1}$				−0.007 (0.008)		0.030*** (0.011)
Percentage of employment among women$_{-1}$				−0.054** (0.025)		−0.006 (0.024)
Percentage of employment in routine occupations$_{-1}$					−0.230*** (0.063)	−0.245*** (0.064)
Average offshorability index of occupations$_{-1}$					0.244 (0.252)	−0.059 (0.237)
Census division dummies	No	No	Yes	Yes	Yes	Yes
	II. 2SLS first stage estimates					
(Δ imports from China to OTH)/worker	0.792*** (0.079)	0.664*** (0.086)	0.652*** (0.090)	0.635*** (0.090)	0.638*** (0.087)	0.631*** (0.087)
R^2	0.54	0.57	0.58	0.58	0.58	0.58

Results

- Rising Chinese imports account for about one-quarter of the aggregate decline in U.S. manufacturing employment.
- A $1,000 per worker increase in import exposure over a decade leads to a **0.60 percentage point** decline in the manufacturing employment share of the working-age population.
- Exposed regions saw higher unemployment, lower labor force participation, and reduced wages.
- Government benefits (e.g., unemployment and disability) increased significantly in more exposed regions.
- The findings highlight significant adjustment costs for local labor markets facing rising low-wage imports.

Suggestive Readings

- Mckenzier David. 2018. *"Rethinking identification under the Bartik Shift-Share Instrument."* World Bank blog
- Byrne, Ieran, Florence Kondylis and John Loeser. 2021. *"Just a little Bartik exposure."* World Bank Blog
- Autor, D., Dorn, D., Hanson, G., & Majlesi, K. 2020. *"Importing political polarization? The electoral consequences of rising trade exposure."* American Economic Review, 110(10), 3139-3183.
- Ferri, Benjamin. 2022. *"Novel Shift-Share Instruments and Their Applications",* No 1053, Boston College Working Papers in Economics, Boston College Department of Economics
- Adão, Rodrigo, Michal Kolesár, and Eduardo Morales. 2019. *"Shift-Share Designs: Theory and Inference."* Quarterly Journal of Economics, 134(4): 1949–2010.
- Bartik, Timothy. 1991. *"Who Benefits from State and Local Economic Development Policies?"* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.
- Blanchard, Olivier, and Lawrence Katz. 1992. *"Regional Evolutions."* Brookings Papers on Economic Activity, 23(1): 1–76.
- Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2022. *"Quasi-Experimental Shift-Share Research Designs."* The Review of Economic Studies, 89(1): 181–213.
- Boustan, Leah, Fernando V Ferreira, Hernan Winkler, and Eric M Zolt. 2013. *"The Effect of Rising Income Inequality on Taxation and Public Expenditures: Evidence from U.S. Municipalities and School Districts, 1970-2000."* The Review of Economics and Statistics, 95(1291).
- Card, David. 2001. *"Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration."* Journal of Labor Economics, 19(1): 22–64.
- Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. *"Bartik Instruments: What, When, Why, and How."* American Economic Review, 110(8): 2586–2624.
- Nunn, Nathan, and Nancy Qian. 2014. *"US Food Aid and Civil Conflict."* American Economic Review, 104(6): 1630–1666.